

CLAIMS

What is claimed is:

1. A method, comprising:
receiving a text string having a plurality of characters; and
performing an unanchored search of a database of a stored patterns
matching one or more characters of the text string using a state machine, wherein
the state machine comprises a ternary content addressable memory (TCAM) and
wherein the performing comprises comparing a state and one of the plurality of
characters with contents of a state field and a character field, respectively, stored
in the TCAM.
2. The method of claim 1, wherein the state is a next state.
3. The method of claim 2, further comprising receiving the next state from an
associated memory.
4. The method of claim 1, wherein the state is an idle state.
5. The method of claim 1, wherein the TCAM implements an Aho-Corasick
algorithm.
6. The method of claim 1, wherein performing the unanchored search
comprises:
traversing the state machine with the text string, wherein the state
machine is traversed with one of the plurality of characters at a time; and
transitioning a state of the state machine based on a stored next state.
7. The method of claim 6, further comprising encoding the next state in a
lookup table.

8. The method of claim 7, wherein the plurality of characters comprises valid and invalid characters and wherein encoding comprises encoding the next state in the state machine if a valid character is received in the text string.
9. The method of claim 8, wherein transitioning further comprises transitioning the state machine to a default state if an invalid character is received in the text string.
10. The method of claim 8, wherein the transitioning is stopped when an invalid character is received.
11. The method of claim 1, wherein performing the unanchored search comprises searching for an exact match of one of the stored patterns.
12. The method of claim 1, wherein performing the unanchored search comprises searching for an inexact match of one of the stored patterns.
13. The method of claim 1, wherein the TCAM has a first width and the text string has a second width greater than the first width of the TCAM.
14. The method of claim 1, wherein each of the plurality of characters has a case, and wherein performing the unanchored search further comprises performing the unanchored search insensitive to the case of one or more of the plurality of characters.
15. The method of claim 14, wherein the text string is encoded in a format having a first plurality of bits, wherein one bit of the first plurality of bits corresponds to the case, wherein the contents of the state field has a second

plurality of bits and wherein performing the search insensitive to the case comprises:

- masking out the one bit corresponding to the case; and
- comparing the first plurality of bits with the second plurality of bits.

16. The method of claim 15, wherein performing the search insensitive to the case further comprises transforming the characters of the text string from a first code to a second code, the second code having a bit unused in the comparing.

17. The method of claim 1, wherein the text string has zero or more wildcard characters, zero or more prefix characters preceding the wildcard characters and zero or more suffix characters succeeding the wildcard characters, and wherein performing the unanchored search comprises:

- searching the database for a first pattern matching the prefix characters;
- and
- searching the database for a second pattern matching the suffix characters.

18. The method of claim 17, wherein performing the unanchored search further comprises creating a count that equals a number of the suffix characters plus a number of the wildcard characters.

19. The method of claim 1, wherein performing the unanchored search comprises:

- comparing, in parallel, N number of the characters with the content of the state field.

20. The method of claim 1, wherein the performing further comprises converging all branches of the state machine, for a given stored pattern, to a single next state when a first number of the characters are matched to the contents of a state field to all state transitions of the branches.
21. The method of claim 20, wherein the single next state is an earlier possible next state for at least one of the branches and wherein the converging comprises transitioning at least one of the branches to the earlier possible next state.
22. The method of claim 19, further comprising:
storing the characters in a first-in-first-out (FIFO) storage element having a plurality of positions;
positioning a read pointer at a first position; and
adjusting the read pointer to a second position by an amount equal to N minus 1.
23. A method, comprising:
receiving a text string having a plurality of characters including a first number of prefix characters, a second number of wildcard characters succeeding the prefix characters, and a third number of suffix characters succeeding the wildcard characters;
performing a first search on a ternary content addressable memory (TCAM) for a first stored pattern matching the prefix characters; and
performing a second search of the TCAM for a second stored pattern matching the suffix characters.
24. The method of claim 23, further comprising creating a count that equals a number of the suffix characters plus a number of the wildcard characters.

25. The method of claim 23, wherein each of the plurality of characters has a case, and wherein the first and second searches are insensitive to the case.
26. The method of claim 23, wherein the TCAM has a first width and the text string has a second width greater than the first width.
27. The method of claim 23, further comprising:
returning a match result when the first stored pattern matches the prefix characters, the second stored pattern matches the suffix characters, and second number of wildcard characters is variable.
28. The method of claim 23, further comprising:
returning a match result when the first stored pattern matches the prefix characters, the second stored pattern matches the suffix characters, and second number of wildcard characters is fixed.
29. The method of claim 28, further comprising:
storing a count value that equals a number of the suffix characters plus the fixed second number of the wildcard characters; and
maintaining a count of incoming characters of the text string after receiving the prefix characters; and
returning the match result when the maintained count is equal to the stored count value.
30. A method, comprising:
receiving a text string having a plurality of characters; and

comparing, in parallel, a plurality of groups of the plurality of characters with states stored in a ternary content addressable memory (TCAM), wherein the groups comprises one or more of the plurality of characters.

31. A method, comprising:

receiving a text string having a plurality of characters; and

performing a search of a database of a stored pattern matching one or more characters of the text string using a state machine, wherein the state machine comprises a ternary content addressable memory (TCAM) and wherein the performing comprises comparing a state and one of the plurality of characters with the contents of a state field and a character field, respectively, stored in the TCAM, wherein each of the plurality of characters has a case, and wherein the search is performed insensitive to the case.

32. The method of claim 31, wherein the text string is encoded in a format having a first plurality of bits, wherein one bit of the first plurality of bits corresponds to the case, wherein the contents of the state field has a second plurality of bits and wherein performing the search insensitive to the case comprises:

masking out the one bit corresponding to the case; and

comparing the first plurality of bits with the second plurality of bits.

33. The method of claim 31, wherein performing the search insensitive to the case further comprises transforming the characters of the text string from a first code to a second code, the second code having a bit unused in the comparing.

34. A method, comprising:

receiving a text string having a plurality of characters;
determining a current state of a state machine and a current character of the text string;
generating a comparand comprising the current state and the current character; and
performing a comparison operation in a ternary content addressable memory (TCAM) using the comparand.

35. The method of claim 34, further comprising:
generating an output of the comparison operation;
accessing an action in an associated memory based on the generated output of the comparison operation; and
executing the action.
36. The method of claim 35, wherein the action is no action.
37. The method of claim 35, further comprising:
accessing a next state in the associated memory based on the generated output;
updating the current state in the comparand with the next state; and
performing another comparison operation in the TCAM using the updated comparand.
38. An apparatus, comprising:
a ternary content addressable memory (TCAM) comprising:
a first field to store a current state of a state machine; and

a second field to store a current character corresponding a character received in a text string.

39. The apparatus of claim 38, wherein the TCAM further comprises:
a third field to store a previous result of the state machine; and
a fourth field to store a count.

40. The apparatus of claim 38, further comprising:
an associated memory coupled to the TCAM, the associated memory comprising:
a first field to store a next state of the state machine;
a second field to store a result of the state machine; and
a third field to store an action for the state machine.

41. The apparatus of claim 40, further comprising:
control circuitry coupled to the associated memory and the TCAM, the control circuitry comprising:
a first-in-first-out (FIFO) storage element to store a plurality of characters of the text string; and
a state register coupled to the FIFO storage element and the TCAM, the state register to store the current character received from the FIFO storage element and transmit the current character to the TCAM.

42. The apparatus of claim 41, wherein the state register is coupled to the associated memory to receive the next state, the state register to store the received next state as the current state.

43. The apparatus of claim 40, wherein the control circuitry further comprises a counter coupled to the state register to increment the count.

44. The apparatus of claim 40, wherein the control circuitry further comprises a rollback circuit coupled to the FIFO storage element to adjust a read pointer positioned at a bit position of the FIFO storage element.

45. A string search apparatus, comprising:
control circuitry to receive a text string having a plurality of characters;
and

a pattern and state database including a ternary content addressable memory (TCAM) coupled to an associated memory, wherein the pattern and state database is operable to perform an unanchored search of the plurality of characters with patterns stored in the TCAM and associated memory.

46. The string search apparatus of claim 45, further comprising a processor coupled to the pattern and state database.

47. The string search apparatus of claim 45, wherein the control circuitry comprises:

a first-in-first-out (FIFO) storage element; and
a register coupled to the FIFO storage element and the TCAM.

48. The apparatus of claim 47, wherein the control circuitry further comprises a rollback circuit coupled to the FIFO storage element.

49. The apparatus of claim 45, wherein the pattern and state database implements an Aho-Corasick algorithm.